PHYSICAL REVIEW D **82,** 112007 (2010)

Measurement of the form factors of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ and determination of the CKM matrix element $|V_{cb}|$

W. Dungel,[8] C. Schwanda,[8] I. Adachi,[6] H. Aihara,[43] T. Aushev,[17,10] T. Aziz,[39] A. M. Bakich,[38] V. Balagura,[10]
E. Barberio,[21] M. Bischofberger,[23] A. Bozek,[27] M. Bračko,[19,11] T. E. Browder,[5] P. Chang,[26] Y. Chao,[26] A. Chen,[24]
P. Chen,[26] B. G. Cheon,[4] R. Chistov,[10] I.-S. Cho,[48] K. Cho,[14] K.-S. Choi,[48] Y. Choi,[37] J. Dalseno,[20,40] A. Drutskoy,[2]
S. Eidelman,[1,31] H. Ha,[15] J. Haba,[6] H. Hayashii,[23] Y. Horii,[42] Y. Hoshi,[41] W.-S. Hou,[26] T. Iijima,[22] K. Inami,[22]
M. Iwabuchi,[48] Y. Iwasaki,[6] N. J. Joshi,[39] T. Julius,[21] J. H. Kang,[48] T. Kawasaki,[29] H. J. Kim,[16] H. O. Kim,[16] J. H. Kim,[14]
M. J. Kim,[16] Y. J. Kim,[3] K. Kinoshita,[2] B. R. Ko,[15] P. Križan,[18,11] P. Krokovny,[6] T. Kuhr,[13] T. Kumita,[44] A. Kuzmin,[1,31]
Y.-J. Kwon,[48] S.-H. Kyeong,[48] M. J. Lee,[36] S.-H. Lee,[15] J. Li,[5] A. Limosani,[21] C. Liu,[35] Y. Liu,[26] D. Liventsev,[10]
R. Louvot,[17] A. Matyja,[27] S. McOnie,[38] H. Miyata,[29] R. Mizuk,[10] G. B. Mohanty,[39] T. Mori,[22] E. Nakano,[32] M. Nakao,[6]
Z. Natkaniec,[27] S. Neubauer,[13] S. Nishida,[6] O. Nitoh,[45] T. Nozaki,[6] T. Ohshima,[22] S. Okuno,[12] S. L. Olsen,[36,5]
G. Pakhlova,[10] C. W. Park,[37] H. Park,[16] H. K. Park,[16] M. Petrič,[11] L. E. Piilonen,[46] M. Prim,[13] M. Röhrken,[13]
M. Rozanska,[27] S. Ryu,[36] H. Sahoo,[5] K. Sakai,[29] Y. Sakai,[6] O. Schneider,[17] A. J. Schwartz,[2] K. Senyo,[22] M. E. Sevior,[21]
J.-G. Shiu,[26] J. B. Singh,[33] P. Smerkol,[11] A. Sokolov,[9] S. Stanič,[30] M. Starič,[11] T. Sumiyoshi,[44] S. Suzuki,[34] S. Tanaka,[6]
G. N. Taylor,[21] Y. Teramoto,[32] K. Trabelsi,[6] S. Uehara,[6] T. Uglov,[10] Y. Unno,[4] S. Uno,[6] P. Urquijo,[21] G. Varner,[5]
K. E. Varvell,[38] K. Vervink,[17] C. H. Wang,[25] P. Wang,[7] M. Watanabe,[29] Y. Watanabe,[12] R. Wedd,[21] K. M. Williams,[46]
E. Won,[15] Y. Yamashita,[28] D. Zander,[13] Z. P. Zhang,[35] P. Zhou,[47] V. Zhulanov,[1,31] and T. Zivko[11]

(The Belle Collaboration)

[1]*Budker Institute of Nuclear Physics, Novosibirsk*
[2]*University of Cincinnati, Cincinnati, Ohio 45221*
[3]*The Graduate University for Advanced Studies, Hayama*
[4]*Hanyang University, Seoul*
[5]*University of Hawaii, Honolulu, Hawaii 96822*
[6]*High Energy Accelerator Research Organization (KEK), Tsukuba*
[7]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing*
[8]*Institute of High Energy Physics, Vienna*
[9]*Institute of High Energy Physics, Protvino*
[10]*Institute for Theoretical and Experimental Physics, Moscow*
[11]*J. Stefan Institute, Ljubljana*
[12]*Kanagawa University, Yokohama*
[13]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe*
[14]*Korea Institute of Science and Technology Information, Daejeon*
[15]*Korea University, Seoul*
[16]*Kyungpook National University, Taegu*
[17]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne*
[18]*Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana*
[19]*University of Maribor, Maribor*
[20]*Max-Planck-Institut für Physik, München*
[21]*University of Melbourne, School of Physics, Victoria 3010*
[22]*Nagoya University, Nagoya*
[23]*Nara Women's University, Nara*
[24]*National Central University, Chung-li*
[25]*National United University, Miao Li*
[26]*Department of Physics, National Taiwan University, Taipei*
[27]*H. Niewodniczanski Institute of Nuclear Physics, Krakow*
[28]*Nippon Dental University, Niigata*
[29]*Niigata University, Niigata*
[30]*University of Nova Gorica, Nova Gorica*
[31]*Novosibirsk State University, Novosibirsk*
[32]*Osaka City University, Osaka*
[33]*Panjab University, Chandigarh*
[34]*Saga University, Saga*
[35]*University of Science and Technology of China, Hefei*
[36]*Seoul National University, Seoul*

1550-7998/2010/82(11)/112007(16)

[37]*Sungkyunkwan University, Suwon*
[38]*School of Physics, University of Sydney, NSW 2006*
[39]*Tata Institute of Fundamental Research, Mumbai*
[40]*Excellence Cluster Universe, Technische Universität München, Garching*
[41]*Tohoku Gakuin University, Tagajo*
[42]*Tohoku University, Sendai*
[43]*Department of Physics, University of Tokyo, Tokyo*
[44]*Tokyo Metropolitan University, Tokyo*
[45]*Tokyo University of Agriculture and Technology, Tokyo*
[46]*IPNAS, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[47]*Wayne State University, Detroit, Michigan 48202*
[48]*Yonsei University, Seoul*

(Received 27 October 2010; published 30 December 2010)

This article describes a determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$ from the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ using 711 fb^{-1} of Belle data collected near the $\Upsilon(4S)$ resonance. We simultaneously measure the product of the form factor normalization $\mathcal{F}(1)$ and the matrix element $|V_{cb}|$ as well as the three parameters ρ^2, $R_1(1)$ and $R_2(1)$, which determine the form factors of this decay in the framework of the heavy quark effective theory. The results, based on about 120 000 reconstructed $B^0 \to D^{*-}\ell^+\nu_\ell$ decays, are $\rho^2 = 1.214 \pm 0.034 \pm 0.009$, $R_1(1) = 1.401 \pm 0.034 \pm 0.018$, $R_2(1) = 0.864 \pm 0.024 \pm 0.008$ and $\mathcal{F}(1)|V_{cb}| = (34.6 \pm 0.2 \pm 1.0) \times 10^{-3}$. The branching fraction of $B^0 \to D^{*-}\ell^+\nu_\ell$ is measured at the same time; we obtain a value of $\mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) = (4.58 \pm 0.03 \pm 0.26)\%$. The errors correspond to the statistical and systematic uncertainties. These results give the most precise determination of the form factor parameters and $\mathcal{F}(1)|V_{cb}|$ to date. In addition, a direct, model-independent determination of the form factor shapes has been carried out.

DOI: 10.1103/PhysRevD.82.112007 PACS numbers: 13.20.He, 12.39.Hg

I. INTRODUCTION

The study of the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ is important for several reasons. The total rate is proportional to the magnitude of the Cabibbo-Kobayashi-Maskawa matrix element V_{cb} [1,2] squared. Experimental investigation of the form factors of the decay can check theoretical models and possibly provide input to more detailed theoretical approaches. In addition, $B^0 \to D^{*-}\ell^+\nu_\ell$ is a major background for charmless semileptonic B decays, such as $B \to \pi\ell\nu$, or semileptonic B decays with large missing energy, including $B \to D^*\tau\nu$. Precise knowledge of the form factors in the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay will thus help to reduce systematic uncertainties in these analyses.

This article is organized as follows: After introducing the theoretical framework for the study of $B^0 \to D^{*-}\ell^+\nu_\ell$ decays in Sec. II, the experimental procedure is presented in detail in Sec. III. This is followed by a discussion of our results and the systematic uncertainties assuming the form factor parameterization of Caprini *et al.* [3] in Sec. IV. Finally, a measurement of the form factor shapes is described in Sec. V.

This paper supersedes our previous result [4], based on a subset of the data used in this analysis.

II. THEORETICAL FRAMEWORK

A. Kinematic variables

The decay $B^0 \to D^{*-}\ell^+\nu_\ell$ [5] proceeds through the tree-level transition shown in Fig. 1. Below we will follow the formulation proposed in reviews [6,7], where the kinematics of this process are fully characterized by four variables as discussed below.

The first is a function of the momenta of the B and D^* mesons, labeled w and defined by

$$w = \frac{P_B \cdot P_{D^*}}{m_B m_{D^*}} = \frac{m_B^2 + m_{D^*}^2 - q^2}{2 m_B m_{D^*}}, \tag{1}$$

where m_B and m_{D^*} are the masses of the B and the D^* mesons (5.279 and 2.010 GeV/c^2, respectively [8]), P_B and P_{D^*} are their four-momenta, and $q^2 = (P_\ell + P_\nu)^2$. In the B rest frame the expression for w reduces to the Lorentz boost $\gamma_{D^*} = E_{D^*}/m_{D^*}$. The ranges of w and q^2 are restricted by the kinematics of the decay, with $q^2_{\min} = 0$ corresponding to

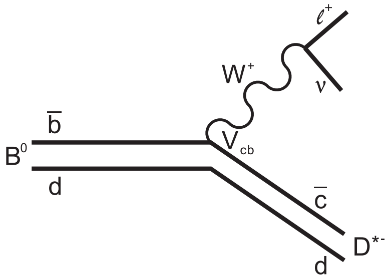

FIG. 1. Quark-level Feynman diagram for the decay $B^0 \to D^{*-}\ell^+\nu_\ell$.

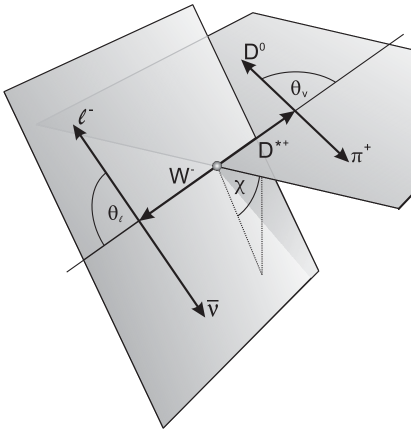

FIG. 2. Definition of the angles θ_ℓ, θ_V and χ for the decay $B^0 \to D^{*-}\ell^+\nu_\ell$, $D^{*-} \to \bar{D}^0\pi_s^-$.

$$w_{\max} = \frac{m_B^2 + m_{D^*}^2}{2m_B m_{D^*}} \approx 1.504, \tag{2}$$

and $w_{\min} = 1$ to

$$q^2_{\max} = (m_B - m_{D^*})^2 \approx 10.69\ \mathrm{GeV}^2. \tag{3}$$

The point $w = 1$ is also referred to as zero recoil.

The remaining three variables are the angles shown in Fig. 2:

(i) θ_ℓ, the angle between the direction of the lepton and the direction opposite the B meson in the virtual W rest frame;

(ii) θ_V, the angle between the direction of the D meson and the direction opposite the B meson in the D^* rest frame;

(iii) χ, the angle between the plane formed by the D^* decay and the plane formed by the W decay, defined in the B meson rest frame.

B. Four-dimensional decay distribution

Three helicity amplitudes, labeled H_+, H_-, and H_0, can be used to describe the Lorentz structure of the $B^0 \to D^{*-}\ell^+\nu_\ell$ decay amplitude. These quantities correspond to the three polarization states of the D^*, two transverse and one longitudinal. When neglecting the lepton mass, i.e., considering only electrons and muons, these amplitudes are expressed in terms of the three functions $h_{A_1}(w)$, $R_1(w)$, and $R_2(w)$ as follows [6]:

$$H_i(w) = m_B \frac{R^*(1 - r^2)(w + 1)}{2\sqrt{1 - 2wr + r^2}} h_{A_1}(w)\tilde{H}_i(w), \tag{4}$$

where

$$\tilde{H}_{\mp} = \frac{\sqrt{1 - 2wr + r^2}\left(1 \pm \sqrt{\frac{w-1}{w+1}}R_1(w)\right)}{1 - r}, \tag{5}$$

$$\tilde{H}_0 = 1 + \frac{(w - 1)(1 - R_2(w))}{1 - r}, \tag{6}$$

with $R^* = (2\sqrt{m_B m_{D^*}})/(m_B + m_{D^*})$ and $r = m_{D^*}/m_B$. The functions $R_1(w)$ and $R_2(w)$ are defined in terms of the axial and vector form factors as

$$A_2(w) = \frac{R_2(w)}{R^{*2}} \frac{2}{w + 1} A_1(w), \tag{7}$$

$$V(w) = \frac{R_1(w)}{R^{*2}} \frac{2}{w + 1} A_1(w). \tag{8}$$

By convention, the function $h_{A_1}(w)$ is defined as

$$h_{A_1}(w) = \frac{1}{R^*} \frac{2}{w + 1} A_1(w). \tag{9}$$

The axial form factor $A_1(w)$ dominates for $w \to 1$. Furthermore, in the limit of infinite b- and c-quark masses, a single form factor describes the decay, the so-called Isgur-Wise function [9,10].

In terms of the three helicity amplitudes, the fully differential decay rate is given by

$$\begin{aligned}
&\frac{d^4\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dw\,d(\cos\theta_\ell)d(\cos\theta_V)d\chi} \\
&\quad = \frac{6m_B m_{D^*}^2}{8(4\pi)^4}\sqrt{w^2 - 1}(1 - 2wr + r^2)G_F^2|V_{cb}|^2 \\
&\quad \times\{(1 - \cos\theta_\ell)^2\sin^2\theta_V H_+^2(w) \\
&\quad + (1 + \cos\theta_\ell)^2\sin^2\theta_V H_-^2(w) + 4\sin^2\theta_\ell\cos^2\theta_V H_0^2(w) \\
&\quad - 2\sin^2\theta_\ell\sin^2\theta_V\cos 2\chi H_+(w)H_-(w) \\
&\quad - 4\sin\theta_\ell(1 - \cos\theta_\ell)\sin\theta_V\cos\theta_V\cos\chi H_+(w)H_0(w) \\
&\quad + 4\sin\theta_\ell(1 + \cos\theta_\ell)\sin\theta_V\cos\theta_V\cos\chi H_-(w)H_0(w)\},
\end{aligned} \tag{10}$$

with $G_F = (1.16637 \pm 0.00001) \times 10^{-5}\hbar c^2\ \mathrm{GeV}^{-2}$. Four one-dimensional decay distributions can be obtained by integrating this decay rate over all but one of the four variables, w, $\cos\theta_\ell$, $\cos\theta_V$, or χ. The differential decay rate as a function of w is

$$\frac{d\Gamma}{dw} = \frac{G_F^2}{48\pi^3} m_{D^*}^3 (m_B - m_{D^*})^2 \mathcal{G}(w)\mathcal{F}^2(w)|V_{cb}|^2, \tag{11}$$

where

$$\begin{aligned}
\mathcal{F}^2(w)\mathcal{G}(w) = {} & h_{A_1}^2(w)\sqrt{w^2 - 1}(w + 1)^2 \\
& \times\left\{2 \times \left[\frac{1 - 2wr + r^2}{(1 - r)^2}\right]\left[1 + R_1(w)^2\frac{w - 1}{w + 1}\right]\right. \\
& \left. + \left[1 + (1 - R_2(w))\frac{w - 1}{1 - r}\right]^2\right\},
\end{aligned}$$

and $\mathcal{G}(w)$ is a known phase space factor,

$$\mathcal{G}(w) = \sqrt{w^2 - 1}(w + 1)^2\left[1 + 4\frac{w}{w + 1}\frac{1 - 2wr + r^2}{(1 - r)^2}\right].$$

A value of the form factor normalization $\mathcal{F}(1) = 1$ is predicted by heavy quark symmetry [6] in the infinite quark-mass limit. Lattice QCD can be utilized to calculate corrections to this limit. The most recent result obtained in unquenched lattice QCD is $\mathcal{F}(1) = 0.921 \pm 0.013 \pm 0.020$ [11].

C. Form factor parameterization

A parameterization of form factors $h_{A_1}(w)$, $R_1(w)$, and $R_2(w)$ can be obtained using heavy quark effective theory (HQET). Perfect heavy quark symmetry implies that $R_1(w) = R_2(w) = 1$, i.e., the form factors A_2 and V are identical for all values of w and differ from A_1 only by a simple kinematic factor. Corrections to this approximation have been calculated in powers of $\Lambda_{\mathrm{QCD}}/m_b$ and the strong coupling constant α_s. Various parameterizations in powers of $(w - 1)$ have been proposed. We adopt the following expressions derived by Caprini, Lellouch, and Neubert [3]:

$$h_{A_1}(w) = h_{A_1}(1)[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3], \tag{12}$$

$$R_1(w) = R_1(1) - 0.12(w - 1) + 0.05(w - 1)^2, \tag{13}$$

$$R_2(w) = R_2(1) + 0.11(w - 1) - 0.06(w - 1)^2, \tag{14}$$

where $z = (\sqrt{w + 1} - \sqrt{2})/(\sqrt{w + 1} + \sqrt{2})$. In addition to the form factor normalization $\mathcal{F}(1) = h_{A_1}(1)$, these expressions contain three free parameters, ρ^2, $R_1(1)$, and $R_2(1)$. The values of these parameters cannot be calculated in a model-independent manner. Instead, they have to be extracted by an analysis of experimental data.

III. EXPERIMENTAL PROCEDURE

A. Data sample and event selection

The data used in this analysis were taken with the Belle detector [12] at the KEKB asymmetric-energy e^+e^- collider [13]. The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight (TOF) scintillation counters, and an electromagnetic calorimeter comprised of CsI(Tl) crystals (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail in Ref. [12]. Two inner detector configurations were used. A 2.0 cm beam pipe and a 3-layer silicon vertex detector were used for the first sample of 152×10^6 $B\bar{B}$ pairs, while a 1.5 cm beam pipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining 620×10^6 $B\bar{B}$ pairs [14].

The data sample consists of 711 fb^{-1} taken at the $\Upsilon(4S)$ resonance, or about 772×10^6 $B\bar{B}$ events. Another 88 fb^{-1} taken at 60 MeV below the resonance are used to estimate the non-$B\bar{B}$ (continuum) background. The off-resonance data is scaled by the integrated on- to off-resonance luminosity ratio corrected for the $1/s$ dependence of the $e^+e^- \to q\bar{q}$ cross section.

This data sample contains events recorded with two different detector setups as well as two different tracking algorithms and large differences in the input files used for Monte Carlo (MC) generation. To ensure that no systematic uncertainty appears due to inadequate consideration of these differences, we separate the data sample into four distinct sets labeled A (141 fb^{-1}), B (274 fb^{-1}), C (189 fb^{-1}) and D (107 fb^{-1}), where the number in parentheses indicates the integrated luminosity corresponding to the individual samples.

Monte Carlo generated samples of $B\bar{B}$ decays equivalent to about 3 times the integrated luminosity are used in this analysis. Monte Carlo simulated events are generated with the EVTGEN program [15], and full detector simulation based on GEANT [16] is applied. QED final state radiation in $B \to X\ell\nu$ decays is added using the PHOTOS package [17].

Hadronic events are selected based on the charged track multiplicity and the visible energy in the calorimeter. The selection is described in detail elsewhere [18]. We also apply a requirement on the ratio of the second to the zeroth Fox-Wolfram moment [19], $R_2 < 0.4$, to reject continuum events.

B. Event reconstruction

Charged tracks are required to originate from the interaction point by applying the following selections on the impact parameters in the $r - \phi$ and z directions: $dr < 2$ cm and $|dz| < 4$ cm, respectively. In addition, we demand at least one associated hit in the SVD detector. For pion and kaon candidates, the Cherenkov light yield from the ACC, the time-of-flight information from TOF, and dE/dx from the CDC are required to be consistent with the appropriate mass hypothesis.

Neutral D meson candidates are reconstructed in the $D^0 \to K^-\pi^+$ decay channel. We fit the charged tracks to a common vertex and reject the D^0 candidate if the χ^2 probability is below 10^{-3}. The reconstructed D^0 mass is required to lie within ± 13.75 MeV$/c^2$ of the nominal D^0 mass of 1.865 GeV$/c^2$ [8], corresponding to about 2.5 times the experimental resolution measured from data.

The D^0 candidate is combined with an additional charged pion (oppositely charged with respect to the kaon candidate) to form a D^{*+} candidate. Because of the kinematics of the D^{*+} decay, the momentum of this pion does not exceed 350 MeV$/c$. It is therefore referred to as the "slow" pion, π_s^+. No impact parameter or hit requirements are applied for π_s. Again, a vertex fit is performed and the same vertex requirement is applied. The invariant mass difference between the D^* and the D candidates, $\Delta m = m_{D^*} - m_{D^0}$, is required to be less than $165\ \mathrm{MeV}/c^2$. This selection is tightened after the background estimation described below. Additional continuum suppression is achieved by requiring that the D^* momentum in the c.m. frame be below $2.45\ \mathrm{GeV}/c$.

Finally, the D^* candidate is combined with an oppositely charged lepton (electron or muon). Electron candidates are identified using the ratio of the energy detected in the ECL to the track momentum, the ECL shower shape, position matching between track and ECL cluster, the energy loss in the CDC, and the response of the ACC counters. Muons are identified based on their penetration range and transverse scattering in the KLM detector. In the momentum region relevant to this analysis, charged leptons are identified with an efficiency of about 90% while the probability to misidentify a pion as an electron (muon) is 0.25% (1.4%) [20,21]. Lepton tracks have to be associated with at least one SVD hit. In the laboratory frame, the momentum of the electron (muon) is required to be greater than $0.30\ \mathrm{GeV}/c$ ($0.60\ \mathrm{GeV}/c$). We also require the lepton momentum in the c.m. frame to be less than $2.4\ \mathrm{GeV}/c$ to reject continuum. More stringent lepton requirements are imposed later in the analysis.

For electron candidates we attempt bremsstrahlung recovery by searching for photons within a cone of 3 degrees around the electron track. If such a photon is found, it is merged with the electron and the sum of the momenta is taken to be the lepton momentum.

C. Background estimation

Because we do not reconstruct the other B meson in the event, the B momentum is *a priori* unknown. However, in the c.m. frame, one can show that the B direction lies on a cone around the $(D^*\ell)$ axis [22] with an opening angle $2\cos\theta_{B,D^*\ell}$ defined by

$$\cos\theta_{B,D^*\ell} = \frac{2E_B^* E_{D^*\ell}^* - m_B^2 - m_{D^*\ell}^2}{2|\vec{p}_B^*||\vec{p}_{D^*\ell}^*|}. \tag{15}$$

In this expression, E_B^* is half of the c.m. energy and $|\vec{p}_B^*|$ is $\sqrt{E_B^{*2} - m_B^2}$. The quantities $E_{D^*\ell}^*$, $\vec{p}_{D^*\ell}^*$ and $m_{D^*\ell}$ are calculated from the reconstructed $D^*\ell$ system.

This cosine is also a powerful discriminator between signal and background: Signal events should lie in the interval $(-1, 1)$, although—due to finite detector resolution—about 5% of the signal is reconstructed outside this interval. The background, on the other hand, does not have this restriction.

The signal lies predominantly in the region defined by $144\ \mathrm{MeV}/c^2 < \Delta m < 147\ \mathrm{MeV}/c^2$ and $p_\ell > 0.8\ \mathrm{GeV}/c$ ($p_\ell > 0.85\ \mathrm{GeV}/c$) for electrons (muons). The region outside these thresholds can be used to estimate the background level.

We therefore perform a fit to the three-dimensional $(\cos\theta_{B,D^*\ell},\ \Delta m,\ p_\ell)$ distributions. The $\cos\theta_{B,D^*\ell}$ range between -10 and 5 is divided into 30 bins. The Δm (p_ℓ) range is divided into five (two) bins, with bin boundaries at 141, 144, 147, 150, 153, $156\ \mathrm{MeV}/c^2$ (0.3, 0.8, $3.5\ \mathrm{GeV}/c$ for electrons and 0.6, 0.85, $3.5\ \mathrm{GeV}/c$ for muons).

The background contained in the final sample has the following six components:

(1) D^{**}: background from $B \to \bar{D}^{**}\ell^+\nu$ decays with $\bar{D}^{**} \to D^*\pi$ or $\bar{D}^{**} \to D\pi$ and from nonresonant $B \to D^*\pi\ell^+\nu$ events, where the lepton has been correctly identified while the D^* candidate may or may not be correctly reconstructed.

(2) correlated background: background from processes other than $B \to \bar{D}^{**}\ell^+\nu$ decays in which both the D^* and the lepton have been correctly reconstructed and originate from the same B meson, e.g., $B^0 \to D^{*-}\tau^+\nu$, $\tau^+ \to \mu^+\nu\bar{\nu}$.

(3) uncorrelated background: the D^* and the lepton have been correctly reconstructed, but they come from different B mesons. In addition, the lepton is not from a $B \to \bar{D}^{**}\ell^+\nu$ decay.

(4) fake lepton: the charged lepton candidate is a misidentified hadron while the D^* candidate may or may not be correctly reconstructed.

(5) fake D^*: the D^* candidate is misreconstructed. The lepton candidate is identified correctly, but it is not from a $B \to \bar{D}^{**}\ell^+\nu$ decay.

(6) continuum: background from $e^+e^- \to q\bar{q}$ ($q = u$, d, s, c processes.

To model the D^{**} component, which consists of a total of four resonant (D_1, D_0^*, D_1', D_2^*) and one nonresonant $D^*\pi\ell\nu$ mode for both neutral and charged B decays, we reweight the branching ratios of each subcomponent to match the values reported by the Particle Data Group [8]. For the resonant parts, only products of branching ratios $\mathcal{B}(B \to D^{**}\ell\nu) \times \mathcal{B}(D^{**} \to D^{(*)}\pi)$ are available and consequently we reweight these products. The shape of the D^{**} momentum distributions is also reweighted in 22 bins of q^2 to match the predictions of the Leibovich- Ligeti-Stewart-Wise (LLSW) model [23,24].

All of the background components are modeled by simulation except for continuum events; these are modeled by off-resonance data. For muon events, the shape of the fake lepton background is corrected by the ratio of the pion fake rate in the experimental data over the same quantity in the Monte Carlo, as measured using $K_S^0 \to \pi^+\pi^-$ decays. The lepton identification efficiency is corrected by the ratio between experimental data and Monte Carlo in $2\gamma \to ee/\mu\mu$ events [20,21]. The $(\cos\theta_{B,D^*\ell},\ \Delta m,\ p_\ell)$ distribution in the data is fitted using the TFRACTIONFITTER algorithm [25] within ROOT [26]. The fit is done separately in each of eight subsamples defined by the experiment range

TABLE I. The signal yield and the signal and background fractions (given in %) for selected events passing the requirements $|\cos\theta_{B,D^*\ell}| < 1$, 144 MeV/$c^2 < \Delta m <$ 147 MeV/c^2 and $p_\ell >$ 0.8 GeV/c ($p_\ell >$ 0.85 GeV/c) for electron (muon) channels.

	A, e	A, μ	B, e	B, μ
Num. candidates	14 802	14 203	29 217	26 894
Signal events	11 609 ± 181	11 139 ± 190	23 029 ± 280	21 002 ± 258
Signal (%)	78.43 ± 1.22	78.43 ± 1.34	78.82 ± 0.96	78.09 ± 0.96
D^{**} (%)	5.63 ± 0.78	4.02 ± 0.86	4.32 ± 0.66	3.90 ± 0.60
Signal correlated (%)	1.07 ± 0.17	1.41 ± 0.25	1.33 ± 0.16	1.71 ± 0.19
Uncorrelated (%)	7.24 ± 0.35	6.01 ± 0.40	7.19 ± 0.31	6.31 ± 0.29
Fake ℓ (%)	0.36 ± 0.17	1.99 ± 0.34	0.50 ± 0.17	2.10 ± 0.23
Fake D^* (%)	2.59 ± 0.12	2.81 ± 0.13	3.07 ± 0.11	2.96 ± 0.10
Continuum (%)	4.68 ± 0.54	5.32 ± 0.59	4.77 ± 0.38	4.93 ± 0.40
	C, e	C, μ	D, e	D, μ
Num. candidates	22 056	20 428	15 871	14 719
Signal events	17 301 ± 240	15 513 ± 235	12 365 ± 189	11 469 ± 205
Signal (%)	78.44 ± 1.09	75.94 ± 1.15	77.91 ± 1.19	77.92 ± 1.39
D^{**} (%)	5.15 ± 0.71	5.22 ± 0.71	4.54 ± 0.72	4.67 ± 0.86
Signal correlated (%)	1.56 ± 0.27	2.07 ± 0.37	2.01 ± 0.26	2.73 ± 0.43
Uncorrelated (%)	6.35 ± 0.35	6.01 ± 0.33	7.33 ± 0.38	6.30 ± 0.40
Fake ℓ (%)	0.75 ± 0.18	2.26 ± 0.28	0.30 ± 0.19	1.68 ± 0.38
Fake D^* (%)	2.86 ± 0.12	2.69 ± 0.11	2.89 ± 0.13	2.80 ± 0.14
Continuum (%)	4.88 ± 0.45	5.81 ± 0.51	5.02 ± 0.53	3.89 ± 0.49

and the lepton type. The results are given in Table I. Figure 3 shows plots of the projections in $\cos\theta_{B,D^*\ell}$ for subsample B.

From the results of the fits we obtain, for each background component j, the overall abundance and the corresponding uncertainty, $N_j^{\mathrm{fit}} \pm \sigma_{N_j}$. On the other hand, MC would have predicted a yield of N_i^{MC} events. By forming the ratio $r_j = N_j^{\mathrm{fit}}/N_j^{\mathrm{MC}}$, we obtain a normalization factor, which allows to correct the MC abundance in each of the background components.

In all fits, the continuum normalization is fixed to the on- to off-resonance luminosity ratio, corrected for the $1/s$ dependence of the $e^+e^- \to q\bar{q}$ cross section. In general, the normalizations obtained by the fit agree well with the MC expectations except for the D^{**} component and the fake ℓ component, which are overestimated in the MC. After the background determination only candidates satisfying the requirements $-1 < \cos\theta_{B,D^*\ell} < 1$, 144 MeV/$c^2 < \Delta m <$ 147 MeV/c^2 and $p_\ell >$ 0.8 GeV/c ($p_\ell >$ 0.85 GeV/c) for electrons (muons) are considered for further analysis.

D. Kinematic variables

To calculate the four kinematic variables defined in Eq. (1) and Fig. 2—w, $\cos\theta_\ell$, $\cos\theta_V$ and χ—which characterize the $B^0 \to D^{*-} l^+ \nu$ decay, we need to determine the B^0 rest frame. The B direction is already known to lie on a cone around the $(D^*\ell)$ axis with an opening angle $2\theta_{B,D^*\ell}$ in the c.m. frame, Eq. (15). To initially determine the B direction, we estimate the c.m. frame momentum vector of the nonsignal B meson by summing the momenta of the remaining particles in the event ($\vec{p}^{\,*}_{\mathrm{incl}}$ [22]) and choose the direction on the cone that minimizes the difference to $-\vec{p}^{\,*}_{\mathrm{incl}}$, Fig. 4.

To obtain $\vec{p}^{\,*}_{\mathrm{incl}}$, we exclude tracks passing far from the interaction point. The minimal requirements depend on the transverse momentum of the track, p_T, and are set to $dr >$ 20 cm (15 cm, 10 cm) or $|dz| >$ 100 cm (50 cm, 20 cm) for a track $p_T <$ 250 MeV/c ($p_T <$ 500 MeV/c, $p_T \geq$ 500 MeV/c). Track candidates that are compatible with a multiply reconstructed track generated by a low-momentum particle spiraling in the central drift chamber are also checked for and only one of the multiple tracks is considered. Unmatched clusters in the barrel region must have an energy greater than 50 MeV. For clusters in the forward (backward) region, the threshold is at 100 MeV (150 MeV). We then compute $\vec{p}_{\mathrm{incl}}$ (in the laboratory frame) by summing the three-momenta of the selected particles,

$$\vec{p}_{\mathrm{incl}} = \sum_i \vec{p}_i, \tag{16}$$

where the index i stands for all particles passing the conditions above, and transform this vector into the c.m. frame. Note that we do not introduce any mass assumption for the charged particles. The energy component of p_{incl} is determined by requiring E^*_{incl} to be $E^*_{\mathrm{beam}} = \sqrt{s}/2$.

With the B^0 rest frame reconstructed in this way, the resolutions in the kinematic variables are found to be about 0.025, 0.049, 0.050, and 13.5° for w, $\cos\theta_\ell$, $\cos\theta_V$ and χ, respectively.

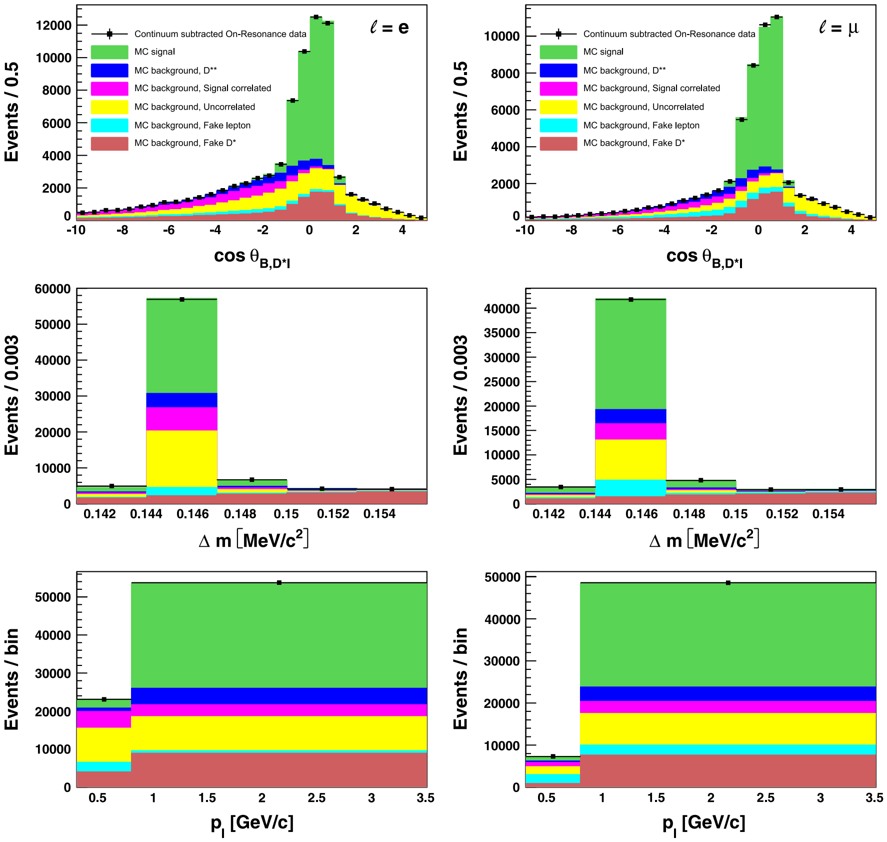

FIG. 3 (color online). Result of the fits to the ($\cos\theta_{B,D^{*0}\ell}$, Δm, p_ℓ) distributions in the e mode (left) and μ mode (right) of the subsample B. The bin boundaries are discussed in the text. The points with error bars are continuum-subtracted on-resonance data. Where not shown, the uncertainties are smaller than the black markers. The histograms are, top to bottom, the signal component, D^{**} background, signal correlated background, uncorrelated background, fake ℓ component and fake D^* component.

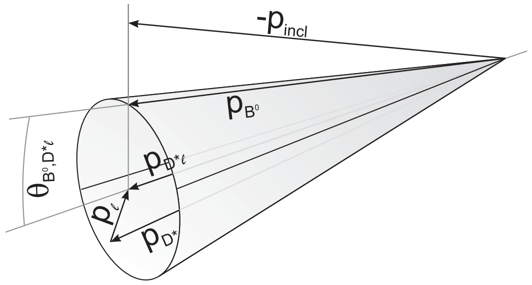

FIG. 4. Reconstruction of the B^0 direction. Refer to the text for details.

IV. ANALYSIS BASED ON THE PARAMETERIZATION OF CAPRINI *ET AL.*

A. Fit procedure

Our main goal is to extract the following quantities: the product of the form factor normalization and $|V_{cb}|$, $\mathcal{F}(1)|V_{cb}|$ (Eq. (11)), and the three parameters ρ^2, $R_1(1)$ and $R_2(1)$ that parameterize the form factors in the HQET framework (Eqs. (12)–(14)). For this, we perform a binned χ^2 fit to the w, $\cos\theta_\ell$, $\cos\theta_V$ and χ distributions over nearly the entire phase space. Instead of an unbinned fit, we fit the one-dimensional projections of w, $\cos\theta_\ell$, $\cos\theta_V$ and χ. This avoids the difficulty of parameterizing the six background components and their correlations in four dimensions. In addition, the one-dimensional projections have sufficient statistics in each bin. However, this approach introduces bin-to-bin correlations that must be accounted for.

The distributions in w, $\cos\theta_\ell$, $\cos\theta_V$ and χ are divided into ten bins of equal width. The kinematically allowed values of w are between 1 and 1.504, but we restrict the fit range to values between 1 and 1.5. In each subsample, there are thus 40 bins to be used in the fit. In the following, we label these bins with a single index i, $i = 1, \ldots, 40$. The bins $i = 1, \ldots, 10$ correspond to the bins of the w distribution, $i = 11, \ldots, 20$ to $\cos\theta_\ell$, $i = 21, \ldots, 30$ to $\cos\theta_V$, and $i = 31, \ldots, 40$ to the χ distribution.

The number of produced events in the bin i, N_i^{th}, is given by

$$N_i^{\mathrm{th}} = N_{B^0}\mathcal{B}(D^{*+} \to D^0\pi^+)\mathcal{B}(D^0 \to K^-\pi^+)\tau_{B^0}\Gamma_i, \tag{17}$$

where N_{B^0} is the number of B^0 mesons in the data sample, $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+)$ are the D^* and D

branching ratios into the final state under consideration [8], τ_{B^0} is the B^0 lifetime [8], and Γ_i is the width obtained by integrating Eq. (10) in the kinematic variable corresponding to i from the lower to the upper bin boundary (the other kinematic variables are integrated over their full range). This integration is numerical in the case of w and analytic for the other variables. The expected number of events $N_i^{\rm exp}$ is related to $N_i^{\rm th}$ as follows:

$$N_i^{\rm exp} = \sum_{j=1}^{40} (R_{ij}\epsilon_j N_j^{\rm th}) + N_i^{\rm bkgrd}. \tag{18}$$

Here, ϵ_j is the probability that an event generated in the bin j is reconstructed and passes all analysis cuts, and R_{ij} is the detector response matrix, i.e., it gives the probability that an event generated in the bin j is observed in the bin i. Both quantities are calculated using MC simulation. R_{ij} takes up a block diagonal form, it describes bin migration between bins of the same kinematic variable. $N_i^{\rm bkgrd}$ is the number of expected background events. For each background component m we apply the normalization r_m, which is obtained from the results of the background estimation as described in Sec. III C. We find

$$N_i^{\rm bkgrd} = \sum_m r_m N_{i,m}^{\rm MCbkgrd}, \tag{19}$$

where $N_{i,m}^{\rm MCbkgrd}$ indicates the number of background events predicted by MC, stemming from component m and falling into bin i.

Next, we calculate the variance σ_i^2 of $N_i^{\rm exp}$. We consider the following contributions: the Poissonian uncertainty in $N_i^{\rm th}$; fluctuations related to the efficiency, estimated by a binomial distribution with N repetitions and known success probability ϵ_i; a similar contribution related to R_{ij} using a multinomial distribution; and the uncertainty in the background contribution $N_i^{\rm bkgrd}$. This yields the following expression for σ_i^2:

$$\begin{aligned}\sigma_i^2 = \sum_{j=1}^{40} \Bigg[& R_{ij}^2 \epsilon_j^2 N_j^{\rm th} + R_{ij}^2 \frac{\epsilon_j(1-\epsilon_j)}{N_{\rm data}} (N_j^{\rm th})^2 \\ & + \frac{R_{ij}(1-R_{ij})}{N'_{\rm data}} \epsilon_j^2 (N_j^{\rm th})^2 + R_{ij}^2 \frac{\epsilon_j(1-\epsilon_j)}{N_{\rm MC}} (N_j^{\rm th})^2 \\ & + \frac{R_{ij}(1-R_{ij})}{N'_{\rm MC}} \epsilon_j^2 (N_j^{\rm th})^2 \Bigg] + \sigma^2(N_i^{\rm bkgrd}).\end{aligned} \tag{20}$$

The first term is the Poissonian uncertainty in $N_i^{\rm th}$. The second and third terms are the binomial and multinomial uncertainties, respectively, related to the finite real data size, where $N_{\rm data}$ ($N'_{\rm data}$) is the total number of decays (the number of reconstructed decays) into the final state under consideration ($K\pi$, e or μ) in the real data. $N_{\rm data}$ is estimated using the total reconstruction efficiency obtained from MC, $N_{i,{\rm data}} = N'_{i,{\rm data}}/\epsilon_i$. The quantities ϵ_i and R_{ij} are calculated from a finite signal MC sample ($N_{\rm MC}$ and $N'_{\rm MC}$); the corresponding uncertainties are estimated by the fourth and fifth terms. Finally, the last term is the background contribution $\sigma^2(N_i^{\rm bkgrd})$, calculated as the sum of the different background component variances. For each background component we estimate its contribution by linear error propagation of the results determined in Sec. III C. We assign a Poisson error to the MC prediction of the number of background events and obtain

$$\sigma^2(N_i^{\rm bkgrd}) = \sum_m (r_m)^2 N_{i,m}^{\rm MCbkgrd}. \tag{21}$$

The uncertainty on the normalization factors r_m is considered in the systematic uncertainty. These variances give the diagonal elements of the covariance matrix C_{ij}.

The bins of different kinematic variables are in general not statistically independent, since the same events are entered into the four histograms. Therefore also the covariance matrix does not take up a diagonal form. In each subsample we calculate the off-diagonal elements of the covariance matrix C_{ij}, $i \neq j$ as $Np_{ij} - Np_i p_j$, where p_{ij} is the relative abundance of bin (i, j) in the two-dimensional histogram obtained by plotting the kinematic variables against each other, p_i is the relative number of entries in the one-dimensional distribution, and N is the size of the sample. Covariances are calculated for the signal and the different background components in the MC samples, and added with appropriate normalizations.

The covariance matrix is inverted numerically within ROOT [26] and, labeling the electron and muon mode in each subsample with the index k, χ^2 functions are calculated,

$$\chi_k^2 = \sum_{i,j} (N_{k,i}^{\rm obs} - N_{k,i}^{\rm exp}) C_{ij}^{-1} (N_{k,j}^{\rm obs} - N_{k,j}^{\rm exp}), \tag{22}$$

where $N_{k,i}^{\rm obs}$ is the number of events observed in bin i in data sample k. We sum these two functions in each subsample and minimize the global χ^2 with MINUIT [27].

We have tested this fit procedure using generic MC data samples. All results are consistent with expectations and show no indication of bias.

B. Investigation of the efficiency of low-momentum tracks

The tracking efficiency of the Belle experiment is reproduced well by MC simulations for tracks with momenta above 200 MeV/c, which we refer to as "high momentum tracks." However, a significant portion of the momentum spectrum of the slow pions emitted in the D^* decay lies below this boundary. For low momenta, the effects of interactions with the detector material such as multiple scattering and energy loss become important and might lead to a deviation between data and MC in the reconstruction efficiency.

We use one half of the reconstructed $B \to D^* \ell \nu$ sample to obtain corrections to the MC reconstruction efficiency in the low-momentum range, measured using real data. The second half is used to perform the analysis with a statistically independent sample. The results of the background estimation shown in Table I are those obtained in the samples used for the analysis. Both of the samples contain about 120 000 signal events.

The sample used to investigate the efficiency of low-momentum tracks is divided into a total of six bins in p_{π_s}. The bin borders of the first five are 50 MeV/c, 100 MeV/c, 125 MeV/c, 150 MeV/c, 175 MeV/c and 200 MeV/c. The region beyond 200 MeV/c defines the sixth bin. By subtracting the background, we obtain an estimate of the signal in data and form the ratio with the signal in MC in each bin, $f_i = N_i^{\mathrm{data}}/N_i^{\mathrm{MC}}$.

The high momentum range is used as normalization, no efficiency correction is applied there. In the lower momentum bins we obtain the ratios $\rho_{\pi_s,i} = f_i/f_{\max}$, which are identical to the ratio of reconstruction efficiencies in the bins i and the high momentum region, $\rho_{\pi_s,i} = \epsilon_i/\epsilon_{\max}$. We calculate this set of ratios for the electron and muon modes and form the weighted average, separately for each of the four subsamples. These values are applied as weights when filling the MC histograms to correct the reconstruction efficiency.

Most systematic uncertainties cancel out in the ratios $\rho_{\pi_s,i}$. Only the uncertainties in the various background components give a small systematic contribution to the uncertainty.

This procedure assumes that the distribution of events in the p_{π_s} spectrum is identical for data and MC. However, one of the aims of the analysis is to measure the form factor parameters that govern this distribution. Therefore, an iterative procedure is adopted: we calculate one set of corrections, apply them and perform the analysis to determine $\mathcal{F}(1)|V_{cb}|$ and the form factor parameters. We then calculate a new set of corrections using these results and repeat the analysis. The changes of the parameters during this iterative procedure are small and vanish after the third iteration. We assign an additional systematic uncertainty to our results based on the stability of the corrections against changes in the form factor parameters. As will be shown in Table III, this is a negligibly small contribution.

C. Results of the fits and investigation of the systematic uncertainties in the subsamples

After applying all analysis cuts and subtracting backgrounds, a total of 123 427 ± 636 signal events are used for the analysis, divided into a total of four experimental subsamples as mentioned above. The result of the fit to these data is shown in Fig. 5 and Table II. The χ^2 per degree of freedom, χ^2/n.d.f, of all fits is good. Table II also gives the χ^2 probabilities or P values, P_{χ^2}.

To estimate the systematic uncertainties in these results, we consider contributions from the following sources: uncertainties in the background component normalizations, uncertainty in the MC tracking efficiency, errors in the world average of $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+)$ as well as in the $\mathcal{B}(B \to D^{**}\ell\nu)$ components [8], uncertainties in the shape of the w distribution of $B \to D^{**}\ell\nu$ events based on the LLSW model [23], uncertainties in the B^0

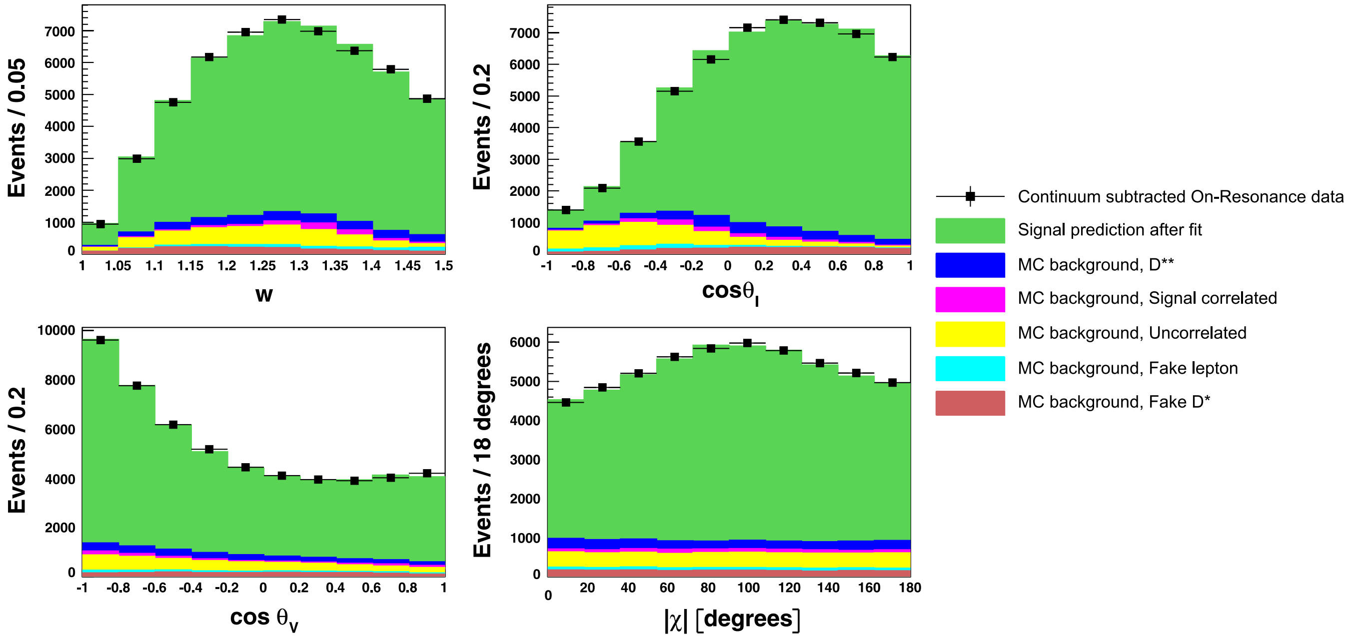

FIG. 5 (color online). Result of the fit of the four kinematic variables in the subsample B. The electron and muon modes are added in this plot. The points with error bars are continuum-subtracted on-resonance data. Where not shown, the uncertainties are smaller than the black markers. The histograms are, top to bottom, the signal component, D^{**} background, signal correlated background, uncorrelated background, fake ℓ component and fake D^* component.

TABLE II. The fit results for the four subsamples. The first two columns show results obtained by investigating only the e or the μ channel, the third column is obtained by minimizing the sum of the χ^2 values calculated for each channel. Where given, the first error is statistical, and the second is systematic.

Sample A	$\ell = e$	$\ell = \mu$	Total sample
ρ^2	$1.248 \pm 0.102 \pm 0.022$	$1.285 \pm 0.114 \pm 0.028$	$1.259 \pm 0.076 \pm 0.019$
$R_1(1)$	$1.317 \pm 0.099 \pm 0.041$	$1.577 \pm 0.131 \pm 0.036$	$1.436 \pm 0.078 \pm 0.030$
$R_2(1)$	$0.804 \pm 0.076 \pm 0.017$	$0.768 \pm 0.093 \pm 0.020$	$0.795 \pm 0.058 \pm 0.015$
$\mathcal{F}(1)\|V_{cb}\| \times 10^3$	$34.8 \pm 0.5 \pm 1.2$	$34.6 \pm 0.6 \pm 1.2$	$34.7 \pm 0.4 \pm 1.2$
χ^2/n.d.f	32.2/36.0	31.6/36.0	70.9/76.0
P_{χ^2}	0.651	0.676	0.643
Sample B	$\ell = e$	$\ell = \mu$	Total sample
ρ^2	$1.169 \pm 0.079 \pm 0.011$	$1.167 \pm 0.088 \pm 0.016$	$1.168 \pm 0.059 \pm 0.011$
$R_1(1)$	$1.411 \pm 0.079 \pm 0.026$	$1.449 \pm 0.090 \pm 0.028$	$1.427 \pm 0.059 \pm 0.022$
$R_2(1)$	$0.902 \pm 0.054 \pm 0.011$	$0.859 \pm 0.061 \pm 0.013$	$0.882 \pm 0.041 \pm 0.010$
$\mathcal{F}(1)\|V_{cb}\| \times 10^3$	$34.4 \pm 0.4 \pm 1.1$	$33.9 \pm 0.4 \pm 1.1$	$34.2 \pm 0.3 \pm 1.1$
χ^2/n.d.f	22.7/36.0	36.5/36.0	60.7/76.0
P_{χ^2}	0.958	0.443	0.900
Sample C	$D^0 \to K\pi$, $\ell = e$	$D^0 \to K\pi$, $\ell = \mu$	Total sample
ρ^2	$1.226 \pm 0.088 \pm 0.011$	$1.262 \pm 0.101 \pm 0.016$	$1.239 \pm 0.066 \pm 0.011$
$R_1(1)$	$1.363 \pm 0.086 \pm 0.026$	$1.480 \pm 0.107 \pm 0.033$	$1.411 \pm 0.066 \pm 0.023$
$R_2(1)$	$0.891 \pm 0.062 \pm 0.012$	$0.851 \pm 0.076 \pm 0.015$	$0.876 \pm 0.048 \pm 0.012$
$\mathcal{F}(1)\|V_{cb}\| \times 10^3$	$34.4 \pm 0.5 \pm 1.1$	$33.9 \pm 0.5 \pm 1.1$	$34.2 \pm 0.3 \pm 1.1$
χ^2/n.d.f	38.6/36.0	38.2/36.0	81.4/76.0
P_{χ^2}	0.352	0.370	0.314
Sample D	$D^0 \to K\pi$, $\ell = e$	$D^0 \to K\pi$, $\ell = \mu$	Total sample
ρ^2	$1.321 \pm 0.102 \pm 0.019$	$1.174 \pm 0.106 \pm 0.020$	$1.247 \pm 0.073 \pm 0.014$
$R_1(1)$	$1.448 \pm 0.109 \pm 0.041$	$1.230 \pm 0.089 \pm 0.031$	$1.330 \pm 0.069 \pm 0.027$
$R_2(1)$	$0.791 \pm 0.081 \pm 0.019$	$0.931 \pm 0.071 \pm 0.015$	$0.864 \pm 0.053 \pm 0.014$
$\mathcal{F}(1)\|V_{cb}\| \times 10^3$	$35.4 \pm 0.6 \pm 1.2$	$35.7 \pm 0.6 \pm 1.2$	$35.6 \pm 0.4 \pm 1.2$
χ^2/n.d.f	25.1/36.0	42.0/36.0	70.1/76.0
P_{χ^2}	0.913	0.226	0.669

lifetime [8], and the uncertainties in the total number of B^0 mesons in the data sample.

To calculate these systematic uncertainties, we consider 300 pseudo experiments in which one of 15 parameters is randomly varied, using a normal distribution. The entire analysis chain is repeated for every pseudo experiment and new fit results are obtained, in total for 4500 variations. One standard deviation in the pseudo-experiment fit results for a given parameter is used as the systematic uncertainty in this parameter.

The parameters varied in the pseudo experiments are as follows:

1. The corrections on the tracking efficiencies for low-momentum tracks are varied within their respective uncertainties. To obtain the most conservative estimate, the uncertainties in different momentum bins are assumed to be fully correlated. Therefore, this component corresponds to a single parameter in the toy MC.
2. The lepton identification efficiencies are varied within their respective uncertainties [20,21].
3. The normalization of the continuum background is not correlated with any of the other backgrounds, it is therefore varied individually within the uncertainty on the on- to off-resonance luminosity ratio, which is 1.0%.
4. The normalizations of the remaining five background components are varied within the uncertainties listed in Table I, while taking into account the correlations found in the background estimation described in Sec. III C.
5. Uncertainties in the composition of the D^{**} component are accounted for by varying each of the components contributing to the D^{**} background within the uncertainty reported by the Particle Data Group [8]. For the resonant modes, this is the uncertainty in the branching fraction products $\mathcal{B}(B \to D^{**}\ell\nu) \times \mathcal{B}(D^{**} \to D^{(*)}\pi)$; for

the nonresonant mode, this is the uncertainty in $\mathcal{B}(B \to D^* \pi \ell \nu)$.

(6) In addition, the shape of the q^2 distributions of the D^{**} components is varied according to the LLSW model [23] and the uncertainties on the model parameters as determined in Ref. [24].

(7) The number of $B^0\bar{B}^0$ events is obtained from the product of the number of $\Upsilon(4S)$ events in the sample with the branching fraction of $\Upsilon(4S)$ to a $B^0\bar{B}^0$ pair. We vary the fraction $f_{+-}/f_{00} = \mathcal{B}(\Upsilon(4S) \to B^+B^-)/\mathcal{B}(\Upsilon(4S) \to B^0\bar{B}^0)$ within its uncertainty [8]. This affects both the overall normalization and the background distributions.

The uncertainties in the reconstruction of the high momentum tracks, the branching ratios $\mathcal{B}(D^{*+} \to D^0\pi^+)$ and $\mathcal{B}(D^0 \to K^-\pi^+)$, the number of $\Upsilon(4S)$ events in the sample, and the B^0 lifetime affect only $\mathcal{F}(1)|V_{cb}|$, not the form factors. Therefore, their uncertainties are considered by analytical error propagation.

D. Averaging the results of the subsamples

To obtain the average of the four subsamples, which have been measured independently, we use the algorithm applied by the Heavy Flavor Averaging Group [28] to obtain the world average for $|V_{cb}|$ from semileptonic B decays. This algorithm combines both the statistical and the systematic uncertainties. The correlations of some of these errors between different samples is considered. For example, the uncertainty on the $D^* \to D^0\pi$ branching fraction will lead to a fully correlated systematic uncertainty in each $B \to D^*\ell\nu$ analysis.

The average is obtained with the MINUIT package [27] by using a χ^2 minimization. Here, N_p gives the total number of fit parameters, in our case $N_p = 4$. When calculating the average of N measurements of the four fit parameters $p_j = \{\mathcal{F}(1)|V_{cb}|, \rho^2, R_1(1), R_2(1)\}$, a total of $4 \times N$ values are available as inputs, which we label as $V_i = \{V_1, \ldots, V_{4N}\}$. In general this number can be labeled as N_i. Each measurement V_i corresponds to one of the parameters p_j, which defines a primitive map $\pi(i)$: $i \to p$. The statistical covariance matrix of each measurement is known, as well as the correlation between the samples. The latter are zero in our case. This information allows one to construct a $4N \times 4N$-dimensional covariance matrix containing the statistical uncertainties and to obtain the statistical part of the χ^2 to be minimized:

$$\chi^2_{\text{stat}} = \sum_i \sum_j (V_i - \bar{V}_{\pi(i)})(C^{-1})_{ij}(V_j - \bar{V}_{\pi(j)}), \qquad (23)$$

where $\bar{V}_k$ indicates the average values of the fit parameters.

The systematic uncertainties are implemented by assuming Gaussian error distributions. The possible bias of input i with respect to the systematic source s can therefore be estimated as $\sigma_{i,s} \times r_s$, where r_s is a normal distributed random number. Its square is by definition distributed according to a χ^2 distribution with one degree of freedom, $r_s^2 \sim \chi^2(1)$. The parameters r_s are floating fit parameters in MINUIT, and in each minimization step the sums

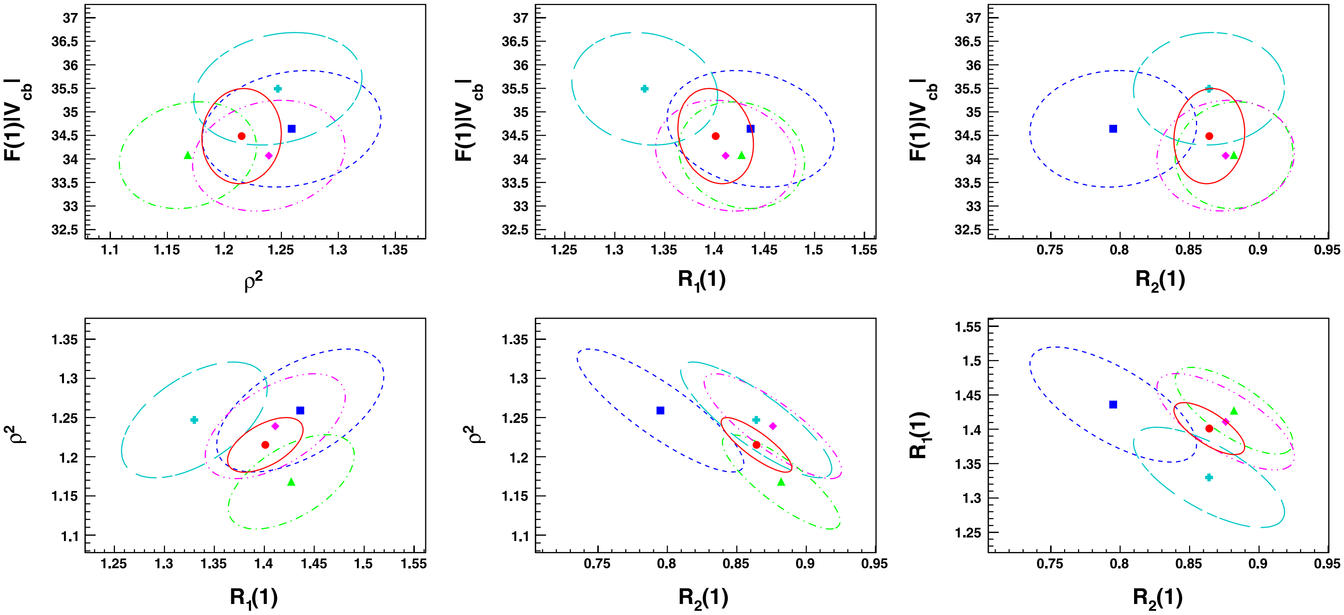

FIG. 6 (color online). Plots of the result of the averaging procedure. Projections in $\mathcal{F}(1)|V_{cb}|$ vs ρ^2 (top left), $\mathcal{F}(1)|V_{cb}|$ vs $R_1(1)$ (top middle), $\mathcal{F}(1)|V_{cb}|$ vs $R_2(1)$ (top right), ρ^2 vs $R_1(1)$ (bottom left), ρ^2 vs $R_2(1)$ (bottom middle) and $R_1(1)$ vs $R_2(1)$ (bottom right) are shown. The red dot (solid line) shows the position (1σ ellipse) of the average, the blue rectangle (dashed line) the position (1σ ellipse) of the subsample A, the green triangle (dash-dotted line) the position (1σ ellipse) of the subsample B, the magenta diamond (dash-double dotted line) the position (1σ ellipse) of the subsample C and the cyan cross (dash-triple dotted line) the position (1σ ellipse) of the subsample D.

TABLE III. The breakup of the systematic uncertainty in the result of the fit to the full sample. The sign + (−) implies whether the fit result moves to larger (smaller) values, if the value of the corresponding systematic parameter is increased.

	ρ^2	$R_1(1)$	$R_2(1)$	$\mathcal{F}(1)\lvert V_{cb}\rvert \times 10^3$	$\mathcal{B}(B^0 \to D^* \ell \nu)$ [%]
Value	1.214	1.401	0.864	34.6	4.58
Statistical error	0.034	0.034	0.024	0.2	0.03
Systematic error	0.009	0.018	0.008	1.0	0.26
Fast track efficiency				−0.78	−0.206
Slow track efficiency	+0.002	+0.003	−0.004	−0.28	−0.059
ρ_{π_s} stability	+0.001	−0.001	+0.000	−0.03	−0.003
LeptonID	+0.002	+0.006	−0.002	−0.38	−0.100
Norm–D^{**}	+0.001	+0.001	−0.001	−0.03	−0.008
Norm–signal corr.	+0.002	−0.003	+0.002	+0.02	+0.006
Norm–uncorr	+0.002	+0.008	−0.003	−0.02	−0.001
Norm–fake ℓ	+0.003	−0.003	−0.001	−0.01	−0.003
Norm–fake D^*	+0.001	−0.001	+0.000	+0.00	+0.003
Norm–continuum	+0.002	+0.002	−0.001	+0.00	−0.003
D^{**} composition	+0.004	+0.009	−0.003	−0.10	−0.025
D^{**} shape	+0.003	+0.005	−0.002	−0.04	−0.011
$N(\Upsilon(4S))$				−0.24	−0.063
f_{+-}/f_{00}	+0.004	−0.009	+0.003	+0.24	+0.062
B^0 lifetime				−0.10	−0.027
$\mathcal{B}(D^* \to D^0 \pi_s)$				−0.13	−0.034
$\mathcal{B}(D^0 \to K\pi)$				−0.22	−0.059

$$\hat{V}_i = V_i + \sum_s \sigma_{i,s} r_s, \quad \forall\ i, \tag{24}$$

are evaluated. If a systematic uncertainty is associated with two different inputs i_1 and i_2, they are both varied by the same fractional systematic uncertainty at the same time. The correlation between systematics is therefore included.

The total χ^2 to be minimized takes the form,

$$\chi^2 = \sum_i \sum_j (\hat{V}_i - \bar{V}_{\pi(i)})(C^{-1})_{ij}(\hat{V}_j - \bar{V}_{\pi(j)}) + \sum_s r_s^2, \tag{25}$$

and is minimized numerically. The number of degrees of freedom are calculated as

$$\text{n.d.f.} = (N_i + N_s) - \underbrace{(N_p + N_s)}_{\text{floated parameters}} = N_i - N_p, \tag{26}$$

which is the same result one obtains in the case without any systematic uncertainties. The minimization is numerically stable and yields both the central values and the total uncertainties of the full four-dimensional average.

Applying this procedure to the four results presented in Table II yields the final result of this analysis. We obtain

TABLE IV. The statistical correlation coefficients of the four parameters in the fit to the full sample.

	$\mathcal{F}(1)\lvert V_{cb}\rvert$	ρ^2	$R_1(1)$	$R_2(1)$
$\mathcal{F}(1)\lvert V_{cb}\rvert$	1.000	0.625	−0.122	−0.206
ρ^2		1.000	0.575	−0.872
$R_1(1)$			1.000	−0.697
$R_2(1)$				1.000

$$\begin{aligned} \mathcal{F}(1)\lvert V_{cb}\rvert &= (34.6 \pm 0.2 \pm 1.0) \times 10^{-3}, \\ \rho^2 &= 1.214 \pm 0.034 \pm 0.009, \\ R_1(1) &= 1.401 \pm 0.034 \pm 0.018, \\ R_2(1) &= 0.864 \pm 0.024 \pm 0.008, \end{aligned} \tag{27}$$

with a $\chi^2/\text{n.d.f} = 14.3/12$ ($P_{\chi^2} = 0.282$). This implies excellent agreement between the results, which can also be seen in the projections of the minimization, shown in Fig. 6. The corresponding branching fraction for the process $B^0 \to D^{*-}\ell^+\nu$ is obtained from the integral of the differential decay width. We obtain

$$\mathcal{B}(B^0 \to D^{*-}\ell^+\nu) = (4.58 \pm 0.03 \pm 0.26)\%. \tag{28}$$

A breakdown of the systematic uncertainties is shown in Table III. The statistical correlation coefficients of the result can be found in Table IV.

V. MODEL-INDEPENDENT DETERMINATION OF HELICITY FUNCTIONS

The angular distributions given in Eq. (10) are determined by the kinematic properties of the decay. However, as discussed in Sec. II C, the expressions of the helicity amplitudes and thus the distribution in the variable w are based on the parameterization scheme proposed by Caprini, Lellouch, and Neubert [3]. In this section, we extract the form factor shape of the longitudinal and the transverse components of Eq. (10) through a fit to the w vs $\cos\theta_V$ distribution. The binning is the same as in the fit approach described above. The contribution from events with $w > 1.5$ is fixed to the small values predicted by the results of the parameterized fit.

A. Fit procedure

From Eq. (10) we can obtain the double differential decay width $d\Gamma/dwd\cos\theta_V$ by integration over $\cos\theta_\ell$ and χ.

If we define

$$F_\Gamma = \frac{G_F^2(m_B - m_{D^*})^2 m_{D^*}^3}{4^3\pi^3} \tag{29}$$

and

$$\begin{aligned} g^{\pm\pm}(w) &= \sqrt{w^2-1}(w+1)^2 h_{A_1}^2(w)|V_{cb}|^2 \\ &\quad\times \frac{1-2wr-r^2}{(1-r)^2}\left\{1 \mp \sqrt{\frac{w-1}{w+1}}R_1(w)\right\}^2, \\ g^{00}(w) &= \sqrt{w^2-1}(w+1)^2 h_{A_1}^2(w)|V_{cb}|^2 \\ &\quad\times\left\{1 + \frac{w-1}{1-r}(1-R_2(w))\right\}^2, \end{aligned} \tag{30}$$

this equation becomes

$$\frac{d^2\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dwd\cos\theta_V} = F_\Gamma[\sin^2\theta_V(g^{++} + g^{--}) + 2\cos^2\theta_V g^{00}]. \tag{31}$$

The quantities g^{kk} correspond to the product of w-dependent parts of the different helicity combinations and kinematic factors. The one-dimensional distribution, as given in Eq. (11), depends only on the sum of these three combinations,

$$\frac{d\Gamma(B^0 \to D^{*-}\ell^+\nu_\ell)}{dw} = \frac{4}{3}F_\Gamma(g^{++} + g^{--} + g^{00}). \tag{32}$$

The bin contents of the two-dimensional histogram in w vs $\cos\theta_V$ can be obtained by integration of Eq. (31) over the corresponding bin area and considering the reconstruction efficiencies and detector response as described in Eq. (18). Each bin content can be given as the linear combination of two linearly independent parts. The integration of the angular distributions is performed analytically, the integration with respect to w defines a set of dimensionless parameters,

$$G_i^{kk} = \int_{w_i}^{w_{i+1}} dw g^{kk}, \tag{33}$$

where $w_j = \{w_1, w_2, \ldots, w_{11}\} = \{1, 1.05, \ldots, 1.5\}$ are the bin boundaries of the 10 bins in w. In addition, we define $g^T = g^{++} + g^{--}$, $G_i^T = G_i^{++} + G_i^{--}$, $g^L = g^{00}$ and $G_i^L = G_i^{00}$.

For the w vs $\cos\theta_V$ distribution we calculate two χ^2 functions

$$\tilde{\chi}_m^2 = \sum_{i=1}^{10}\sum_{j=1}^{10}\frac{N_{m,ij}^{2D,\mathrm{obs}} - N_{m,ij}^{2D,\mathrm{exp}}}{\sigma^2_{N_{m,ij}^{2D,\mathrm{exp}}}}, \tag{34}$$

which depends only on the parameters G_i^T and G_i^L. Here the index m denotes the two considered decay channels (e and μ), N^{obs} gives the number of events observed in on-resonance data, N^{exp} the number of expected events, as defined in Eq. (18), and $\sigma_{N^{\mathrm{exp}}}$ the uncertainty in the expected number of events, as given in Eq. (20). Again we form the sum of these two χ^2 functions, and minimize this expression numerically using MINUIT [27].

We have tested this fit procedure using generic MC data samples. All results are consistent with expectations and show no indication of bias.

B. Results

We investigated the largest of the four subsamples—corresponding to about 274 fb^{-1} of data recorded on the $\Upsilon(4S)$ resonance—to extract the helicity shapes. Tables V and VI give the results of the fits, where the systematic errors quoted in these tables originate from the same sources given in the breakdown in Table III. Many of the dominant systematic uncertainties, such as the track reconstruction or the lepton ID uncertainty, are correlated between different bins. The χ^2 of the fit is statistically consistent with the number of degrees of freedom, we obtain $\chi^2/\mathrm{n.d.f} = 175.8/179$, $P_{\chi^2} = 55.4\%$. The results are shown in Fig. 7 and Tables V and VI, compared to the values obtained using the parameters given in Sec. IV D.

Finally we compute a numerical measure of the agreement between the result obtained from this two-dimensional fit with the functions predicted by the parameterization of Caprini, Lellouch, and Neubert [3]. Comparing with the values given in Table II is naturally problematic, since the underlying sample is identical in both fits. Therefore we compare the extracted shapes to the parameterization using the world average reported by the Heavy Flavor Averaging Group [28] in the summer of 2010 and form a simple χ^2,

$$\chi^2 = \sum_i\sum_j(G_i^{\mathrm{param}} - G_i^{\mathrm{fit}})(C_{\mathrm{stat}}^{-1})_{ij}(G_j^{\mathrm{param}} - G_j^{\mathrm{fit}}) \tag{35}$$

TABLE V. Results obtained for G_i^T (dimensionless), compared to the central values obtained from the parameterized fit.

	$D^0 \to K\pi$, $\ell = e$	$D^0 \to K\pi$, $\ell = \mu$
G_1^T	$(1.187 \pm 0.148 \pm 0.075) \times 10^{-4}$	$(0.982 \pm 0.142 \pm 0.067) \times 10^{-4}$
G_2^T	$(1.514 \pm 0.112 \pm 0.102) \times 10^{-4}$	$(1.239 \pm 0.104 \pm 0.081) \times 10^{-4}$
G_3^T	$(1.594 \pm 0.116 \pm 0.104) \times 10^{-4}$	$(1.685 \pm 0.116 \pm 0.115) \times 10^{-4}$
G_4^T	$(1.809 \pm 0.122 \pm 0.113) \times 10^{-4}$	$(1.760 \pm 0.128 \pm 0.115) \times 10^{-4}$
G_5^T	$(1.649 \pm 0.136 \pm 0.105) \times 10^{-4}$	$(1.484 \pm 0.141 \pm 0.102) \times 10^{-4}$
G_6^T	$(1.511 \pm 0.145 \pm 0.097) \times 10^{-4}$	$(1.572 \pm 0.165 \pm 0.104) \times 10^{-4}$
G_7^T	$(1.135 \pm 0.156 \pm 0.069) \times 10^{-4}$	$(0.974 \pm 0.161 \pm 0.063) \times 10^{-4}$
G_8^T	$(0.933 \pm 0.159 \pm 0.060) \times 10^{-4}$	$(1.072 \pm 0.156 \pm 0.070) \times 10^{-4}$
G_9^T	$(0.631 \pm 0.163 \pm 0.038) \times 10^{-4}$	$(0.571 \pm 0.151 \pm 0.036) \times 10^{-4}$
G_{10}^T	$(0.254 \pm 0.141 \pm 0.025) \times 10^{-4}$	$(0.324 \pm 0.122 \pm 0.038) \times 10^{-4}$
	Fit to total sample	Central value of parametrized fit
G_1^T	$(1.088 \pm 0.102 \pm 0.069) \times 10^{-4}$	0.919×10^{-4}
G_2^T	$(1.388 \pm 0.077 \pm 0.092) \times 10^{-4}$	1.505×10^{-4}
G_3^T	$(1.637 \pm 0.081 \pm 0.108) \times 10^{-4}$	1.706×10^{-4}
G_4^T	$(1.794 \pm 0.085 \pm 0.113) \times 10^{-4}$	1.733×10^{-4}
G_5^T	$(1.547 \pm 0.097 \pm 0.101) \times 10^{-4}$	1.642×10^{-4}
G_6^T	$(1.552 \pm 0.109 \pm 0.100) \times 10^{-4}$	1.466×10^{-4}
G_7^T	$(1.054 \pm 0.111 \pm 0.065) \times 10^{-4}$	1.222×10^{-4}
G_8^T	$(1.000 \pm 0.110 \pm 0.064) \times 10^{-4}$	0.926×10^{-4}
G_9^T	$(0.600 \pm 0.110 \pm 0.035) \times 10^{-4}$	0.589×10^{-4}
G_{10}^T	$(0.297 \pm 0.091 \pm 0.029) \times 10^{-4}$	0.221×10^{-4}

TABLE VI. Results obtained for G_i^L (dimensionless), compared to the central values obtained from the parameterized fit.

	$D^0 \to K\pi$, $\ell = e$	$D^0 \to K\pi$, $\ell = \mu$
G_1^L	$(0.405 \pm 0.083 \pm 0.027) \times 10^{-4}$	$(0.283 \pm 0.090 \pm 0.019) \times 10^{-4}$
G_2^L	$(0.878 \pm 0.096 \pm 0.054) \times 10^{-4}$	$(0.935 \pm 0.099 \pm 0.061) \times 10^{-4}$
G_3^L	$(1.102 \pm 0.109 \pm 0.068) \times 10^{-4}$	$(1.124 \pm 0.112 \pm 0.073) \times 10^{-4}$
G_4^L	$(1.230 \pm 0.128 \pm 0.077) \times 10^{-4}$	$(1.123 \pm 0.133 \pm 0.071) \times 10^{-4}$
G_5^L	$(1.232 \pm 0.137 \pm 0.074) \times 10^{-4}$	$(1.787 \pm 0.151 \pm 0.112) \times 10^{-4}$
G_6^L	$(1.479 \pm 0.149 \pm 0.087) \times 10^{-4}$	$(1.281 \pm 0.159 \pm 0.078) \times 10^{-4}$
G_7^L	$(1.426 \pm 0.152 \pm 0.086) \times 10^{-4}$	$(1.727 \pm 0.171 \pm 0.106) \times 10^{-4}$
G_8^L	$(1.458 \pm 0.154 \pm 0.083) \times 10^{-4}$	$(1.107 \pm 0.165 \pm 0.067) \times 10^{-4}$
G_9^L	$(1.678 \pm 0.146 \pm 0.100) \times 10^{-4}$	$(1.794 \pm 0.154 \pm 0.111) \times 10^{-4}$
G_{10}^L	$(1.592 \pm 0.125 \pm 0.097) \times 10^{-4}$	$(1.527 \pm 0.122 \pm 0.100) \times 10^{-4}$
	Fit to total sample	Central value of parametrized fit
G_1^L	$(0.361 \pm 0.060 \pm 0.025) \times 10^{-4}$	0.480×10^{-4}
G_2^L	$(0.895 \pm 0.069 \pm 0.056) \times 10^{-4}$	0.849×10^{-4}
G_3^L	$(1.118 \pm 0.078 \pm 0.070) \times 10^{-4}$	1.061×10^{-4}
G_4^L	$(1.164 \pm 0.091 \pm 0.073) \times 10^{-4}$	1.213×10^{-4}
G_5^L	$(1.532 \pm 0.102 \pm 0.094) \times 10^{-4}$	1.328×10^{-4}
G_6^L	$(1.372 \pm 0.110 \pm 0.082) \times 10^{-4}$	1.417×10^{-4}
G_7^L	$(1.568 \pm 0.114 \pm 0.095) \times 10^{-4}$	1.486×10^{-4}
G_8^L	$(1.298 \pm 0.112 \pm 0.076) \times 10^{-4}$	1.540×10^{-4}
G_9^L	$(1.730 \pm 0.105 \pm 0.104) \times 10^{-4}$	1.580×10^{-4}
G_{10}^L	$(1.557 \pm 0.085 \pm 0.095) \times 10^{-4}$	1.608×10^{-4}

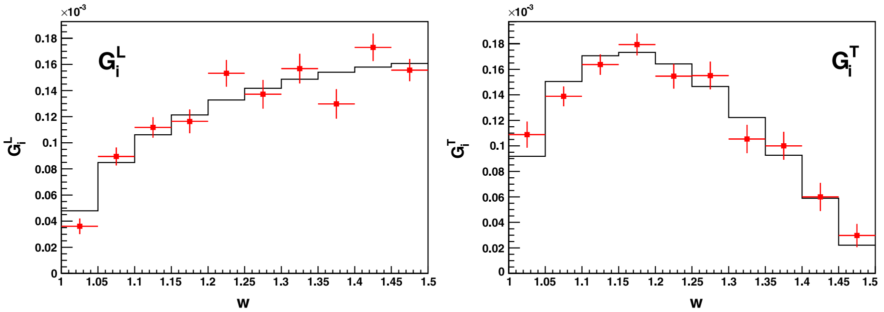

FIG. 7 (color online). Results of the fit of the helicity functions (red crosses) compared to the prediction obtained by using the parameters obtained by using the parameterization prescription by Caprini, Lellouch and Neubert (solid black line). The left plot shows the results for G_i^L, the right one for G_i^T. Only the statistical errors are shown.

where $G_i^{\rm param}$ are the values obtained using the parameterization by Caprini, Lellouch, and Neubert, $G_i^{\rm fit}$ are the values obtained by the unparameterized fit described above and $C_{\rm stat}$ is the covariance matrix between the fit parameters, which is also a result of the numerical minimization. Evaluating this expression yields a $\chi^2/{\rm n.d.f} = 29.3/20$, $P_{\chi^2} = 8.3\%$. This implies satisfactory agreement between the extracted shapes of the longitudinal and transverse components and the world average parameters.

VI. SUMMARY AND DISCUSSION

Using 711 $\rm fb^{-1}$ of data collected by the Belle experiment, we have analyzed approximately 120 000 $B^0 \to D^{*-}\ell^+\nu_\ell$ decays. A fit to four kinematic variables fully characterizing these decays yields measurements of the product of the form factor normalization and $|V_{cb}|$, $\mathcal{F}(1)|V_{cb}|$, and of the parameters ρ^2, $R_1(1)$ and $R_2(1)$ that enter the HQET form factor parameterization of this decay. We obtain

$$\begin{aligned} \mathcal{F}(1)|V_{cb}| &= (34.6 \pm 0.2 \pm 1.0) \times 10^{-3}, \\ \rho^2 &= 1.214 \pm 0.034 \pm 0.009, \\ R_1(1) &= 1.401 \pm 0.034 \pm 0.018, \\ R_2(1) &= 0.864 \pm 0.024 \pm 0.008, \\ \mathcal{B}(B^0 \to D^{*-}\ell^+\nu_\ell) &= (4.58 \pm 0.03 \pm 0.26)\%. \end{aligned} \tag{36}$$

For all these measurements, the first error is the statistical uncertainty and the second is the systematic uncertainty. Using a recent lattice QCD result, $\mathcal{F}(1) = 0.921 \pm 0.013 \pm 0.020$ [11], we obtain the following value of $|V_{cb}|$:

$$|V_{cb}| = 37.5 \pm 0.2 \pm 1.1 \pm 1.0, \tag{37}$$

where the third error is due to the theoretical uncertainty on $\mathcal{F}(1)$. Our results (36) are compatible with the recent measurements of these quantities by the *BABAR* experiment [29–31] as well as with results reported by the ALEPH [32], CLEO [33], DELPHI [34,35] and OPAL [36] experiments. This paper supersedes our previous result [4], based on a subset of the data used in this analysis. The results presented here give the most precise determination of the form factor parameters and $\mathcal{F}(1)|V_{cb}|$ to date.

A direct, model-independent determination of the form factor shapes has also been carried out and is in good agreement with the HQET-based form factor parameterization by Caprini *et al.* [3].

ACKNOWLEDGMENTS

We thank the KEKB group for the excellent operation of the accelerator, the KEK cryogenics group for the efficient operation of the solenoid, and the KEK computer group and the National Institute of Informatics for valuable computing and SINET3 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council and the Australian Department of Industry, Innovation, Science and Research; the National Natural Science Foundation of China under Contract Nos. 10575109, 10775142, 10875115 and 10825524; the Ministry of Education, Youth and Sports of the Czech Republic under Contract Nos. LA10033 and MSM0021620859; the Department of Science and Technology of India; the BK21 and WCU program of the Ministry Education Science and Technology, National Research Foundation of Korea, and NSDC of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area ("New Development of Flavor Physics"), and from JSPS for Creative Scientific Research ("Evolution of Tau-lepton Physics").

[1] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[2] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).

[3] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).

[4] K. Abe *et al.* (Belle Collaboration), Phys. Lett. B **526**, 247 (2002).

[5] Throughout this note charge-conjugate decay modes are implied.

[6] M. Neubert, Phys. Rep. **245**, 259 (1994).

[7] J. D. Richman and P. R. Burchat, Rev. Mod. Phys. **67**, 893 (1995).

[8] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).

[9] N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989).

[10] N. Isgur and M. B. Wise, Phys. Lett. B **237**, 527 (1990).

[11] C. Bernard *et al.*, Phys. Rev. D **79**, 014506 (2009).

[12] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).

[13] S. Kurokawa, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume.

[14] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A **560**, 1 (2006).

[15] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[16] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, CERN Report No. CERN-DD/EE/84-1.

[17] E. Barberio and Z. Was, Comput. Phys. Commun. **79**, 291 (1994).

[18] K. Abe *et al.* (Belle Collaboration), Phys. Rev. D **64**, 072001 (2001).

[19] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[20] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 490 (2002).

[21] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **491**, 69 (2002).

[22] Quantities evaluated in the c.m. frame are marked with an asterisk.

[23] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D **57**, 308 (1998).

[24] P. Urquijo *et al.* (Belle Collaboration), Phys. Rev. D **75**, 032001 (2007).

[25] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).

[26] R. Brun and F. Rademakers, Nucl. Instrum. Methods Phys. Res., Sect. A **389**, 81 (1997).

[27] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).

[28] D. Asner *et al.* (The Heavy Flavor Averaging Group), arXiv:1010.1589.

[29] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **77**, 032002 (2008).

[30] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **100**, 231803 (2008).

[31] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **79**, 012002 (2009).

[32] D. Buskulic *et al.* (ALEPH Collaboration), Phys. Lett. B **395**, 373 (1997).

[33] R. A. Briere *et al.* (CLEO Collaboration), Phys. Rev. Lett. **89**, 081803 (2002).

[34] P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **510**, 55 (2001).

[35] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **33**, 213 (2004).

[36] G. Abbiendi *et al.* (OPAL Collaboration), Phys. Lett. B **482**, 15 (2000).